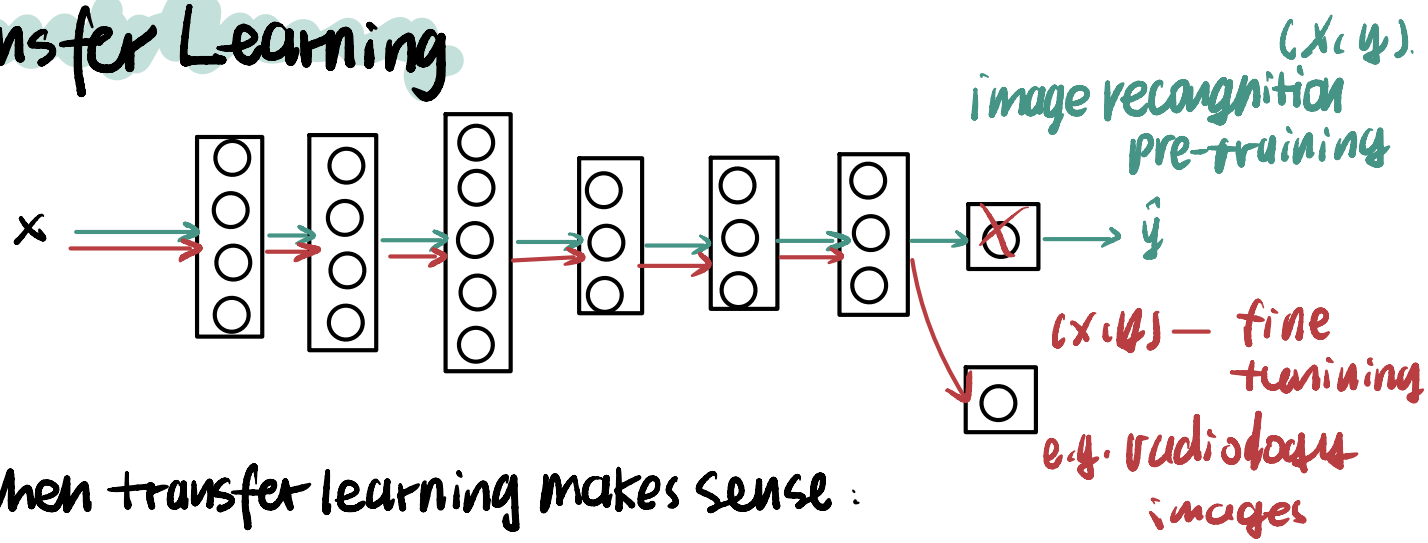
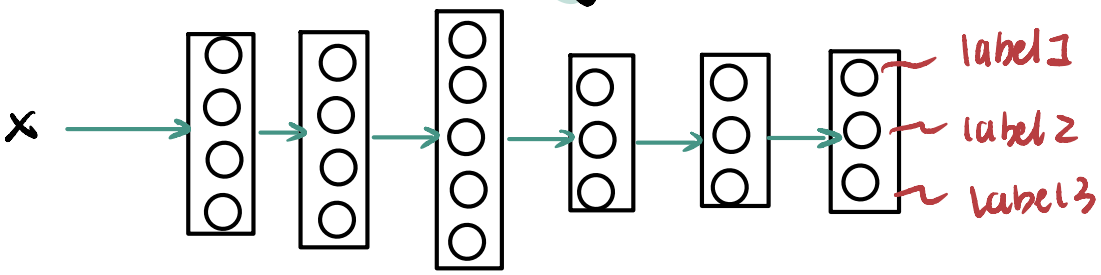
Transfer Learning



When transfer learning makes sense:

- input A and B have the same input X
- you have a lot more data for task A than task B
- lower level feature from A could be helpful for Learning B.

Multi-task learning



loss : $\hat{y}^{(i)}$ $(3, 1)$ $\longrightarrow$ $\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{3} \mathcal{L}(\hat{y}_j^{(i)}, y_j^{(i)})$

When multi-task learning makes sense:

- Training on set of tasks that could benefit from having shared lower-level feature.
- Usually: Amount of data you have for each task is quite similar.
- Can train a big enough neural network to do well on all the tasks